

RECEIVED

JUL 11 2003

TECH CENTER 1600/2

RECEIVED
NOV 12 2003
TECH CENTER 1600/2

1600

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/840,008A

DATE: 07/08/2003

TIME: 08:00:00

Input Set : A:\88802211.app

Output Set: N:\CRF4\07082003\I840008A.raw

3 <110> APPLICANT: EVANS, RONALD M.
5 <120> TITLE OF INVENTION: XENOBIOTIC COMPOUND MODULATED EXPRESSION SYSTEMS AND
6 USES THEREFOR
8 <130> FILE REFERENCE: 088802-5211
10 <140> CURRENT APPLICATION NUMBER: 09/840,008A
11 <141> CURRENT FILING DATE: 2001-04-20
13 <150> PRIOR APPLICATION NUMBER: 09/227,718
14 <151> PRIOR FILING DATE: 1999-01-08
16 <150> PRIOR APPLICATION NUMBER: 09/458,366
17 <151> PRIOR FILING DATE: 1999-12-09
19 <150> PRIOR APPLICATION NUMBER: 09/005,286
20 <151> PRIOR FILING DATE: 1998-01-09
22 <160> NUMBER OF SEQ ID NOS: 44
24 <170> SOFTWARE: PatentIn Ver. 2.1
26 <210> SEQ ID NO: 1
27 <211> LENGTH: 2068
28 <212> TYPE: DNA
29 <213> ORGANISM: Homo sapiens
31 <220> FEATURE:
32 <221> NAME/KEY: CDS
33 <222> LOCATION: (583)..(1887)
35 <400> SEQUENCE: 1

36 ggcacgagga gatctagggtt caaattaatg ttgcccttag tggtaaagga cagagaccct 60
38 cagactgatg aaatgcgctc agaattactt agacaaagcg gatatttgcc actctcttcc 120
40 ccttttcctg tgtttttgta gtgaagagac ctgaaagaaa aaagtaggga gaacataatg 180
42 agaacaaata cggtaatctc ttcaatttgct agttcaagt ctggacttgg gacttaggag 240
44 gggcaatgga gccgcttagt gcctacatct gacttggact gaaatatagg tgagagacaa 300
46 gattgtctca tatccgggga aatcataacc tatgactagg acgggaagag gaagcactgc 360
48 ctttacttca gtgggaatct cggcctcagc ctgcaagcca agtggttcaca gtgagaaaag 420
50 caagagaata agctaatact cctgtcctga acaaggcagc ggctccttgg taaagctact 480
52 ccttgatcga tcctttgcac cggattgttc aaagtggacc ccaggggaga agtcggagca 540
54 aagaacttac caccaagcag tccaagaggc ccagaagcaa ac ctg gag gtg aga 594
55 Met Glu Val Arg
56 1
58 ccc aaa gaa agc tgg aac cat gct gac ttt gta cac tgt gag gac aca 642
59 Pro Lys Glu Ser Trp Asn His Ala Asp Phe Val His Cys Glu Asp Thr
60 5 10 15 20
62 gag tct gtt cct gga aag ccc agt gtc aac gca gat gag gaa gtc gga 690
63 Glu Ser Val Pro Gly Lys Pro Ser Val Asn Ala Asp Glu Glu Val Gly
64 25 30 35
66 ggt ccc caa atc tgc cgt gta tgt ggg gac aag gcc act ggc tat cac 738
67 Gly Pro Gln Ile Cys Arg Val Cys Gly Asp Lys Ala Thr Gly Tyr His
68 40 45 50

ENTERED

RECEIVED

NOV 06 2003

RAW SEQUENCE LISTING

DATE: 07/08/2003

PATENT APPLICATION: US/09/840,008A

TIME: 08:00:44

Input Set : A:\88802211.app

Output Set: N:\CRF4\07082003\I840008A.raw

```

70 ttc aat gtc atg aca tgt gaa gga tgc aag ggc ttt ttc agg agg gcc 786
71 Phe Asn Val Met Thr Cys Glu Gly Cys Lys Gly Phe Phe Arg Arg Ala
72          55          60          65
74 atg aaa cgc aac gcc cgg ctg agg tgc ccc ttc cgg aag ggc gcc tgc 834
75 Met Lys Arg Asn Ala Arg Leu Arg Cys Pro Phe Arg Lys Gly Ala Cys
76      70      75      80
78 gag atc acc cgg aag acc cgg cga cag tgc cag gcc tgc cgc ctg cgc 882
79 Glu Ile Thr Arg Lys Thr Arg Arg Gln Cys Gln Ala Cys Arg Leu Arg
80  85      90      95      100
82 aag tgc ctg gag agc ggc atg aag aag gag atg atc atg tcc gac gag 930
83 Lys Cys Leu Glu Ser Gly Met Lys Lys Glu Met Ile Met Ser Asp Glu
84          105          110          115
86 gcc gtg gag gag agg cgg gcc ttg atc aag cgg aag aaa agt gaa cgg 978
87 Ala Val Glu Glu Arg Arg Ala Leu Ile Lys Arg Lys Lys Ser Glu Arg
88          120          125          130
90 aca ggg act cag cca ctg gga gtg cag ggg ctg aca gag gag cag cgg 1026
91 Thr Gly Thr Gln Pro Leu Gly Val Gln Gly Leu Thr Glu Glu Gln Arg
92          135          140          145
94 atg atg atc agg gag ctg atg gac gct cag atg aaa acc ttt gac act 1074
95 Met Met Ile Arg Glu Leu Met Asp Ala Gln Met Lys Thr Phe Asp Thr
96      150      155      160
98 acc ttc tcc cat ttc aag aat ttc cgg ctg cca ggg gtg ctt agc agt 1122
99 Thr Phe Ser His Phe Lys Asn Phe Arg Leu Pro Gly Val Leu Ser Ser
100 165      170      175      180
102 ggc tgc gag ttg cca gag tct ctg cag gcc cca tcg agg gaa gaa gct 1170
103 Gly Cys Glu Leu Pro Glu Ser Leu Gln Ala Pro Ser Arg Glu Glu Ala
104          185          190          195
106 gcc aag tgg agc cag gtc cgg aaa gat ctg tgc tct ttg aag gtc tct 1218
107 Ala Lys Trp Ser Gln Val Arg Lys Asp Leu Cys Ser Leu Lys Val Ser
108          200          205          210
110 ctg cag ctg cgg ggg gag gat ggc agt gtc tgg aac tac aaa ccc cca 1266
111 Leu Gln Leu Arg Gly Glu Asp Gly Ser Val Trp Asn Tyr Lys Pro Pro
112          215          220          225
114 gcc gac agt ggc ggg aaa gag atc ttc tcc ctg ctg ccc cac atg gct 1314
115 Ala Asp Ser Gly Gly Lys Glu Ile Phe Ser Leu Leu Pro His Met Ala
116      230      235      240
118 gac atg tca acc tac atg ttc aaa ggc atc atc agc ttt gcc aaa gtc 1362
119 Asp Met Ser Thr Tyr Met Phe Lys Gly Ile Ile Ser Phe Ala Lys Val
120 245      250      255      260
122 atc tcc tac ttc agg gac ttg ccc atc gag gac cag atc tcc ctg ctg 1410
123 Ile Ser Tyr Phe Arg Asp Leu Pro Ile Glu Asp Gln Ile Ser Leu Leu
124          265          270          275
126 aag ggg gcc gct ttc gag ctg tgt caa ctg aga ttc aac aca gtg ttc 1458
127 Lys Gly Ala Ala Phe Glu Leu Cys Gln Leu Arg Phe Asn Thr Val Phe
128          280          285          290
130 aac gcg gag act gga acc tgg gag tgt ggc cgg ctg tcc tac tgc ttg 1506
131 Asn Ala Glu Thr Gly Thr Trp Glu Cys Gly Arg Leu Ser Tyr Cys Leu
132          295          300          305
134 gaa gac act gca ggt ggc ttc cag caa ctt cta ctg gag ccc atg ctg 1554

```

RAW SEQUENCE LISTING

DATE: 07/08/2003

PATENT APPLICATION: US/09/840,008A

TIME: 08:00:44

Input Set : A:\88802211.app

Output Set: N:\CRF4\07082003\I840008A.raw

```

135 Glu Asp Thr Ala Gly Gly Phe Gln Gln Leu Leu Leu Glu Pro Met Leu
136      310                      315                      320
138 aaa ttc cac tac atg ctg aag aag ctg cag ctg cat gag gag gag tat 1602
139 Lys Phe His Tyr Met Leu Lys Lys Leu Gln Leu His Glu Glu Glu Tyr
140 325                      330                      335                      340
142 gtg ctg atg cag gcc atc tcc ctc ttc tcc cca gac cgc cca ggt gtg 1650
143 Val Leu Met Gln Ala Ile Ser Leu Phe Ser Pro Asp Arg Pro Gly Val
144      345                      350                      355
146 ctg cag cac cgc gtg gtg gac cag ctg cag gag caa ttc gcc att act 1698
147 Leu Gln His Arg Val Val Asp Gln Leu Gln Glu Gln Phe Ala Ile Thr
148      360                      365                      370
150 ctg aag tcc tac att gaa tgc aat cgg ccc cag cct gct cat agg ttc 1746
151 Leu Lys Ser Tyr Ile Glu Cys Asn Arg Pro Gln Pro Ala His Arg Phe
152      375                      380                      385
154 ttg ttc ctg aag atc atg gct atg ctc acc gag ctc cgc agc atc aat 1794
155 Leu Phe Leu Lys Ile Met Ala Met Leu Thr Glu Leu Arg Ser Ile Asn
156      390                      395                      400
158 gct cag cac acc cag cgg ctg ctg cgc atc cag gac ata cac ccc ttt 1842
159 Ala Gln His Thr Gln Arg Leu Leu Arg Ile Gln Asp Ile His Pro Phe
160 405                      410                      415                      420
162 gct acg ccc ctc atg cag gag ttg ttc ggt atc aca ggt agc tga 1887
163 Ala Thr Pro Leu Met Gln Glu Leu Phe Gly Ile Thr Gly Ser
164      425                      430
166 gtggctgtcc ttgggtgaca cctccgagag gtagttagac ccagagccct ctgagtcgcc 1947
168 actcccgggc caagacagat ggacactgcc aagagccgac aatgccctgc tggcctgtct 2007
170 ccctagggaa ttccctgctat gacagctggc tagcattcct caggaaggac atgggggtgcc 2067
172 c 2068
175 <210> SEQ ID NO: 2
176 <211> LENGTH: 434
177 <212> TYPE: PRT
178 <213> ORGANISM: Homo sapiens
180 <400> SEQUENCE: 2
181 Met Glu Val Arg Pro Lys Glu Ser Trp Asn His Ala Asp Phe Val His
182 1 5 10 15
184 Cys Glu Asp Thr Glu Ser Val Pro Gly Lys Pro Ser Val Asn Ala Asp
185 20 25 30
187 Glu Glu Val Gly Gly Pro Gln Ile Cys Arg Val Cys Gly Asp Lys Ala
188 35 40 45
190 Thr Gly Tyr His Phe Asn Val Met Thr Cys Glu Gly Cys Lys Gly Phe
191 50 55 60
193 Phe Arg Arg Ala Met Lys Arg Asn Ala Arg Leu Arg Cys Pro Phe Arg
194 65 70 75 80
196 Lys Gly Ala Cys Glu Ile Thr Arg Lys Thr Arg Arg Gln Cys Gln Ala
197 85 90 95
199 Cys Arg Leu Arg Lys Cys Leu Glu Ser Gly Met Lys Lys Glu Met Ile
200 100 105 110
202 Met Ser Asp Glu Ala Val Glu Glu Arg Arg Ala Leu Ile Lys Arg Lys
203 115 120 125
205 Lys Ser Glu Arg Thr Gly Thr Gln Pro Leu Gly Val Gln Gly Leu Thr

```

RECEIVED

NOV 06 2003

OFFICE OF PETITIONS

RAW SEQUENCE LISTING

DATE: 07/08/2003

PATENT APPLICATION: US/09/840,008A

TIME: 08:00:44

Input Set : A:\88802211.app

Output Set: N:\CRF4\07082003\I840008A.raw

```

206      130      135      140
208 Glu Glu Gln Arg Met Met Ile Arg Glu Leu Met Asp Ala Gln Met Lys
209 145      150      155      160
211 Thr Phe Asp Thr Thr Phe Ser His Phe Lys Asn Phe Arg Leu Pro Gly
212      165      170      175
214 Val Leu Ser Ser Gly Cys Glu Leu Pro Glu Ser Leu Gln Ala Pro Ser
215      180      185      190
217 Arg Glu Glu Ala Ala Lys Trp Ser Gln Val Arg Lys Asp Leu Cys Ser
218      195      200      205
220 Leu Lys Val Ser Leu Gln Leu Arg Gly Glu Asp Gly Ser Val Trp Asn
221      210      215      220
223 Tyr Lys Pro Pro Ala Asp Ser Gly Gly Lys Glu Ile Phe Ser Leu Leu
224 225      230      235      240
226 Pro His Met Ala Asp Met Ser Thr Tyr Met Phe Lys Gly Ile Ile Ser
227      245      250      255
229 Phe Ala Lys Val Ile Ser Tyr Phe Arg Asp Leu Pro Ile Glu Asp Gln
230      260      265      270
232 Ile Ser Leu Leu Lys Gly Ala Ala Phe Glu Leu Cys Gln Leu Arg Phe
233      275      280      285
235 Asn Thr Val Phe Asn Ala Glu Thr Gly Thr Trp Glu Cys Gly Arg Leu
236      290      295      300
238 Ser Tyr Cys Leu Glu Asp Thr Ala Gly Gly Phe Gln Gln Leu Leu Leu
239 305      310      315      320
241 Glu Pro Met Leu Lys Phe His Tyr Met Leu Lys Lys Leu Gln Leu His
242      325      330      335
244 Glu Glu Glu Tyr Val Leu Met Gln Ala Ile Ser Leu Phe Ser Pro Asp
245      340      345      350
247 Arg Pro Gly Val Leu Gln His Arg Val Val Asp Gln Leu Gln Glu Gln
248      355      360      365
250 Phe Ala Ile Thr Leu Lys Ser Tyr Ile Glu Cys Asn Arg Pro Gln Pro
251      370      375      380
253 Ala His Arg Phe Leu Phe Leu Lys Ile Met Ala Met Leu Thr Glu Leu
254 385      390      395      400
256 Arg Ser Ile Asn Ala Gln His Thr Gln Arg Leu Leu Arg Ile Gln Asp
257      405      410      415
259 Ile His Pro Phe Ala Thr Pro Leu Met Gln Glu Leu Phe Gly Ile Thr
260      420      425      430
262 Gly Ser
265 <210> SEQ ID NO: 3
266 <211> LENGTH: 25
267 <212> TYPE: DNA
268 <213> ORGANISM: Artificial Sequence
270 <220> FEATURE:
271 <223> OTHER INFORMATION: Description of Artificial Sequence: Putative SXR
272      response element from the steroid hydroxylase,
273      rCYP3A1
275 <400> SEQUENCE: 3
276 tagacagttc atgaagttca tctac
279 <210> SEQ ID NO: 4

```

25

RAW SEQUENCE LISTING

DATE: 07/08/2003

PATENT APPLICATION: US/09/840,008A

TIME: 08:00:44

Input Set : A:\88802211.app

Output Set: N:\CRF4\07082003\I840008A.raw

```

280 <211> LENGTH: 25
281 <212> TYPE: DNA
282 <213> ORGANISM: Artificial Sequence
284 <220> FEATURE:
285 <223> OTHER INFORMATION: Description of Artificial Sequence: Putative SXR
286     response element from the steroid hydroxylase,
287     rCYP3A2
289 <400> SEQUENCE: 4
290 taagcagttc ataaagttca tctac                                25
293 <210> SEQ ID NO: 5
294 <211> LENGTH: 25
295 <212> TYPE: DNA
296 <213> ORGANISM: Artificial Sequence
298 <220> FEATURE:
299 <223> OTHER INFORMATION: Description of Artificial Sequence: Putative SXR
300     response element from the steroid hydroxylase,
301     rUGT1A6
303 <400> SEQUENCE: 5
304 actgtagttc ataaagttca catgg                                25
307 <210> SEQ ID NO: 6
308 <211> LENGTH: 26
309 <212> TYPE: DNA
310 <213> ORGANISM: Artificial Sequence
312 <220> FEATURE:
313 <223> OTHER INFORMATION: Description of Artificial Sequence: Putative SXR
314     response element from the steroid hydroxylase,
315     rbCYP2C1
317 <400> SEQUENCE: 6
318 caatcagttc aacagggttc accaat                                26
321 <210> SEQ ID NO: 7
322 <211> LENGTH: 33
323 <212> TYPE: DNA
324 <213> ORGANISM: Artificial Sequence
326 <220> FEATURE:
327 <223> OTHER INFORMATION: Description of Artificial Sequence: Putative SXR
328     response element from the steroid hydroxylase,
329     rP450R
331 <400> SEQUENCE: 7
332 cacagtgag ctgaggccag cagcaggtcg aaa                        33
335 <210> SEQ ID NO: 8
336 <211> LENGTH: 27
337 <212> TYPE: DNA
338 <213> ORGANISM: Artificial Sequence
340 <220> FEATURE:
341 <223> OTHER INFORMATION: Description of Artificial Sequence: Putative SXR
342     response element from the steroid hydroxylase,
343     rCYP2A1
345 <400> SEQUENCE: 8
346 gtgcaggttc aactggaggt caacatg                                27

```

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 07/08/2003
PATENT APPLICATION: US/09/840,008A TIME: 08:00:45

Input Set : A:\88802211.app
Output Set: N:\CRF4\07082003\I840008A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:22; N Pos. 7,8,9,10,11
Seq#:23; N Pos. 7,8,9,10,11,12
Seq#:44; N Pos. 7,8,9,10,11